



USGS Community for Data Integration Data Management Working Group: Science Data Catalog, Open Data Discussion

November 10, 2014

U.S. Department of the Interior U.S. Geological Survey

Goals

- 1. Background on Administration Open Data requirements, DOI implementation, etc.
- Overview of USGS Science Data Catalog (SDC), pathways to data.doi.gov, data.gov, and the Geospatial Platform
 - a. Providing records to SDC, data.doi.gov, data.gov and the geoplatform
- 3. Inventory and status of NGDA and A16 datasets (NGDA and A16 leads)--tracking
- 4. Next steps (all)





Background & Drivers...

1. On May 9, 2013, the White House released the Executive Order, "Making Open and Machine Readable the New Default for Government." This Order, built upon an earlier interagency memorandum released by the OSTP, was accompanied by an OMB Policy requiring:

Policy requiring:

Inventory of Data

Public Data Listing

- data.doi.gov
- data.gov
- USGS Science Data Management activities
 - Community for Data Integration
 - Office of Science Quality and Integrity policy revisions
 - Moving toward a "science data life cycle approach" in USGS
 - Implementation of USGS Science Strategy Plans







Background & Drivers cont...

Helping address Big Earth Data Initiative (BEDI) actions:

- Making large, complex datasets accessible, machine actionable
- Treating these datasets to improve interoperability, accessibility
- Coordinating across DOI, and with NASA, NOAA





Stakeholders & Project Goals...



Project Goals:

- Least impact to all
- Self-service model
- Broad involvement
- Increase Program visibility
- Proper attribution
- Incremental updates
- Improve Data Mgmt.
- Effective links to related efforts
- Reuse of Standards
- Leverage existing systems, efforts
- Move towards actionable data



What is the USGS Science Data Catalog (SDC)?

- 1. Single point for USGS to provide its metadata to data.doi.gov, data.gov, OMB, geoplatform etc.
- 2. Organizes USGS data find, get, use (Actionable)
- 3. Meets Open Data requirements for USGS
- 4. Provides metadata retrieval, visualization, download, & links back to original providers





What is the USGS SDC (cont)?

- Allows data stewards control of what data to publish and review metadata for release – through the "SDC Dashboard"
- 2. Requires metadata content prior to data release--FGDC CSDGM, no broken links (See http://sunlightfoundation.com/blog/2014/10/30/federal-open-data-audit-defense-downright-dismal-interior-immense-yet-imperfect/ for an example of the perils of broken links...)
- 3. Focus on scientific data

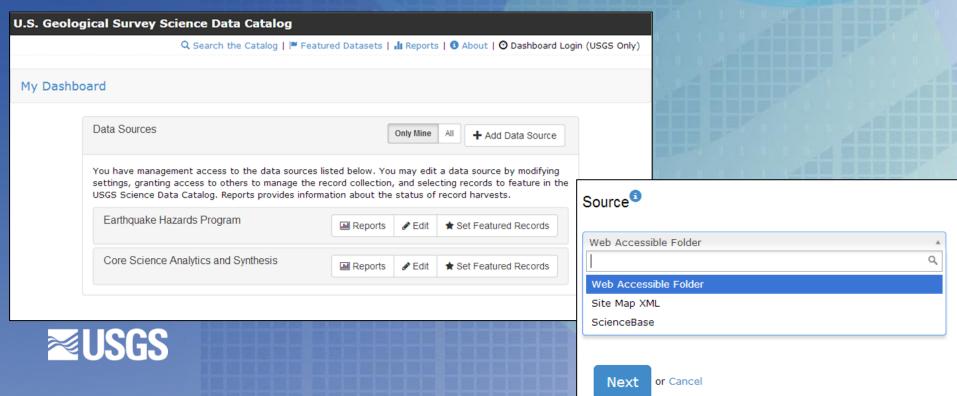




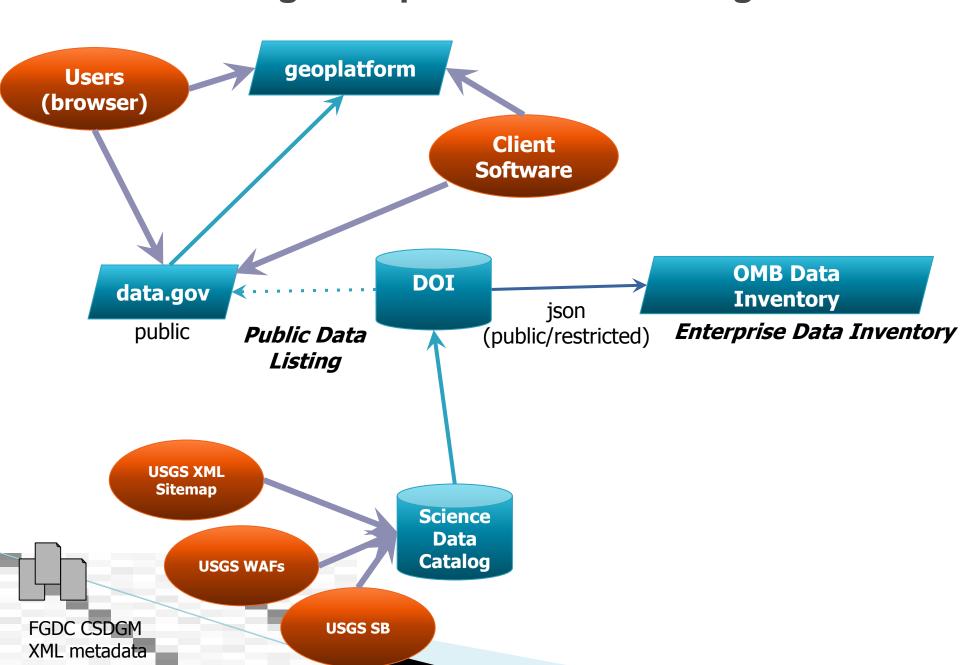
Science Data Catalog (SDC) Dashboard

Allows managers and data stewards to:

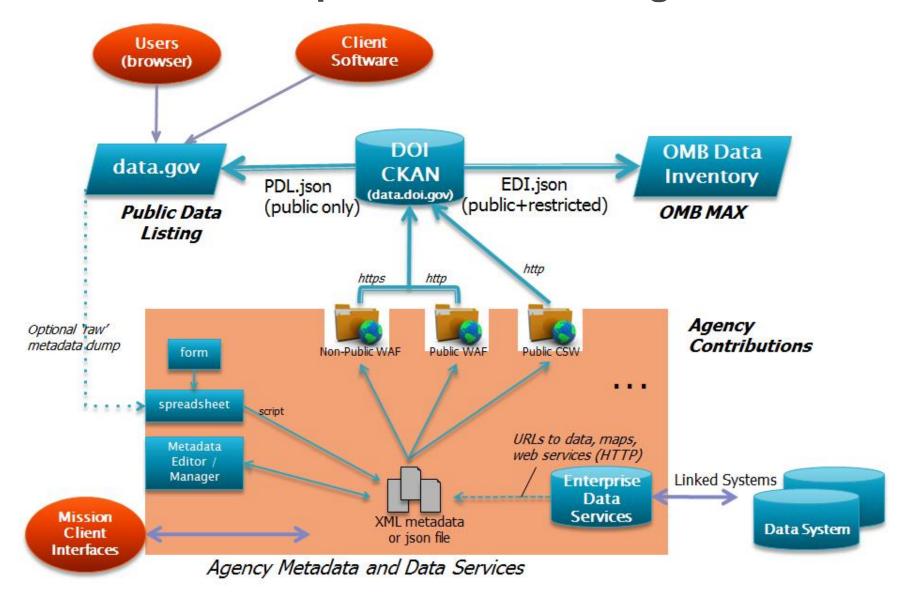
- Register/Define/Maintain metadata sources
- Obtain statistics on metadata indexing, completeness, and errors



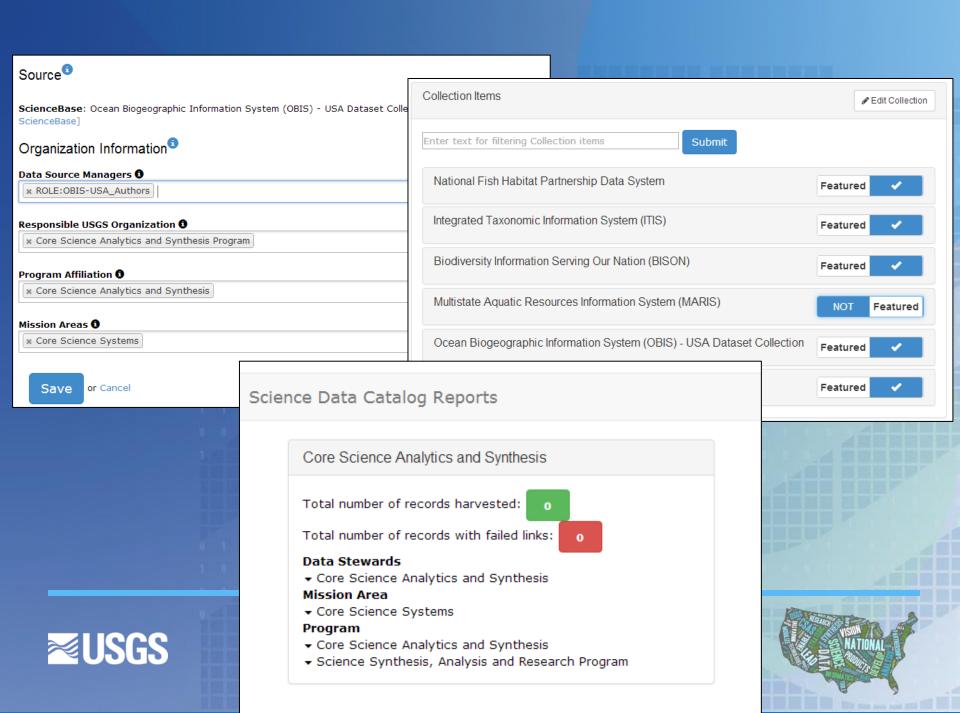
Data Publishing for Open Data and data.gov



Data Release for Open Data and data.gov



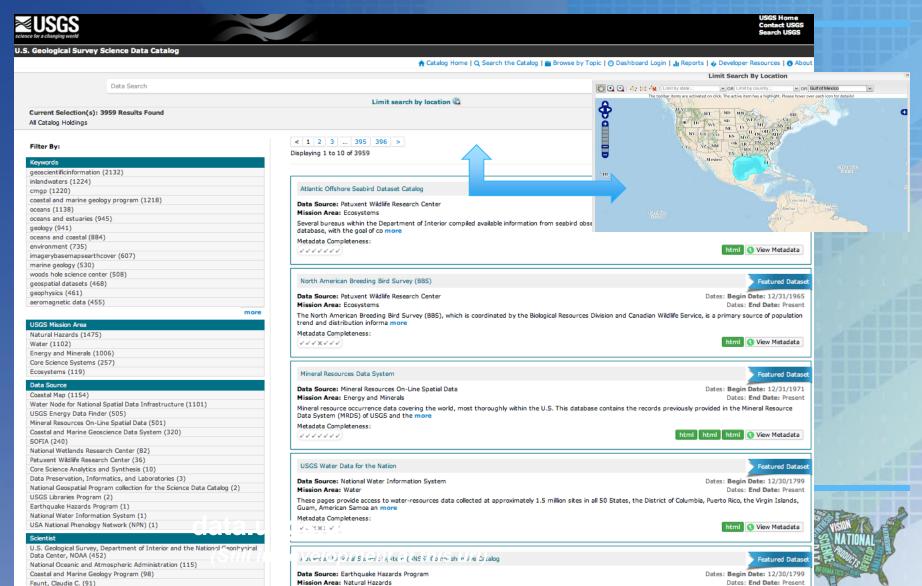




USGS Science Data Catalog-data.usgs.gov

VeeAnn A. Cross (80)

Robert Morton (79)



(e.g. moment tensor solutions, macroseismic more

The ANSS Comprehensive Catalog (ComCat) contains earthquake source parameters (e.g. hypocenters, magnitudes, phase picks and amplitudes) and other products

DOI -- Data.doi.gov







Data.gov



Search Data/Gov

Q

DATA TOPICS- IMPACT APPLICATIONS DEVELOPERS CONTACT

ATA CATALOG		#/Datasets	Organizations	•				
♣/Organizations/U.S. G	eological Survey, Department							
SIICCC	⚠ Datasets							
science for a changing world			Order by:					
U.S. Geological	Search	Q,	Select an option					
Survey, Department of the Interior	10,220 datasets found							
http://www.usgs.gov/ reed more	Datasets ordered by Popular							
Followers Members	USGS National Geologic Map Database Collection 8							
2 25	The National Geologic Map Database (NGMDB) is a Congressionally mandated national archive of geoscience maps, reports, and stratigraphic information. According to the Geologic							
10.2k	HTML SITML							
10.2K	Map Service Showing Geology, Oil and Gas Fields, and Geological Provinces of Iran							
Dataset Type Clear All	The geology data set for this map includes ancs, polygons, and labels that outline and describe the general geologic age and type of bedrock of Iran. The geologic provinces data							
geospatial (10155)	WARS ENGINEEY PITML ENGISSE HITML WARS I move in dataset							
non-geospatial (65)	Oil and Car Euglaration and Braduction in the	a State of Winele C	Thomas ar					
Show More Detayet Type	Oil and Gas Exploration and Production in the State of Illinois Shown as Quarter-Mile Cells 8							
Tags Coar All	A cells polygon feature class was created by the U. S. Geological Survey (USGS) to illustrate the degree of exploration, type of production, and distribution of production in							
O CONT	ZIP ZIP MITHL ZIP ZIP MITHL 1-more in dataset:							
oceans (7746)	USCS OUT A STATE OF THE PARTY O							
marine geology (7407)	USGS Oil and Gas Assessment Database 4 The U.S. Geological Survey (USGS) has assessed undiscovered oil and natural gas potential for 212							
earth science (7250)	geological provinces throughout the world as part of the		Porter that the KTK					
usa (6088)	skapefile, earl personal geodutabases, coverages, tab-del							
north pacific ocean (3628)	shapefile, esri parsonal geodalabases, Coverages, lab-det	Rolled, comma separati	nd, portakte čorument	torn				
Show More Tage	Map Service Showing Geology, Oil and Gas F	ields, and Geologic	Provinces					
400000	of Europe including Turkey 🎳							

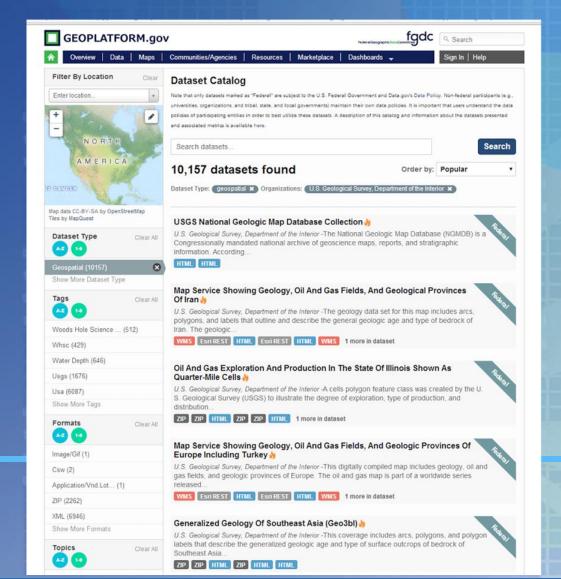
This digitally compiled map includes geology, oil and gas fields, and geologic provinces of Europe.

The oil and gas map is part of a worldwide series released on CD-ROM by the...





Geospatial Platform

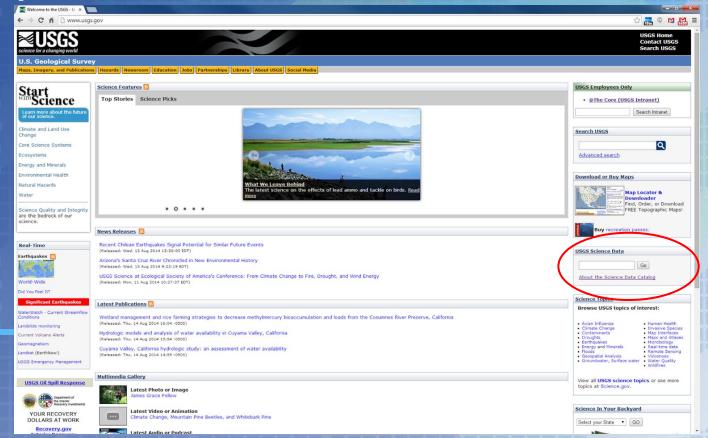






USGS Science Data Catalog Integration with New USGS Web...

- Partnering with OCAP and AEI
- Drive "data" piece of new USGS Web
- Minimizes duplicate effort





SDC Metadata Data Stewards & Sources...

Source Name	Organizational Unit	Mission Area	Point of Contact	Catalog Status	System-level Record or Individual Records?	"Featured Dataset"?	Records Passing Validation in the Catalog
Mineral Resources Data (MRData)	Minerals Program	Energy & Minerals and Environmental Heath	Peter Schweitzer	In Catalog	вотн	Yes, for system-level record	501
William Nesources Data (William)	Williotals F Togram	Hodii	1 otor ocriwolizor	iii Catalog	DOTT	Yes, for 2 records: Breeding	301
						Bird Survey & Atlantic Seabird	
Patuxent Wildlife Research Center	Patuxent Wildlife Research Center	Ecosystems	Mark Wimer	In Catalog	Individual Records	Database	35
US Topographic Maps	National Geospatial Program	Core Science Systems	Rick Brown/Brian Fox	In Catalog	System-level Record	Yes	1
Historical Topographic Maps	National Geospatial Program	Core Science Systems	Rick Brown/Brian Fox	In Catalog	System-level Record	Yes	1
mistorical Topographic maps	National Cooperative Geologic Mapping	Core ocience dystems	Nick Blown/Blian Fox	iii Catalog	System-level Necolu	165	
National Geologic Map Database	Program	Core Science Systems	Dave Soller	Not Yet Started	System-level Record	Yes	not yet started
•		•	John				
A d A	5000	011	Faundeen/Wayne		0		
Landsat	EROS	Climate & Land Use Change	Miller Harley Benz, ANSS	In Process	System-level Record	Yes	in progress
			Technical Manager,				
			(benz@usgs.gov);				
			Cecily Wolfe, ANSS				
			Coordinator,				
Advanced National Seismic Systems (ANSS)	Earthquakes Program	Natural Hazards	(cwolfe@usgs.gov);	In Catalog	System-level Record	Yes	1
USGS Water Data for the Nation	National Water Information System	Water	Dorrie Gellenbeck	In Catalog	System-level Record	Yes	1
National Hydrography Dataset	National Geospatial Program	Core Science Systems	Jeff Simley (sp?)	Not Yet Started	System-level Record	Yes	not yet started
Did You See It? NAWQA Data Warehouse	Landslides Program CIDA	Natural Hazards Water	Rex Baum? Nate Booth	Not Yet Started Not Yet Started	System-level Record System-level Record	Yes Yes	not yet started not yet started
NAVVQA Data Warehouse	CIDA	vvater	Gerald "Stinger"	Not ret Started	System-level Record	res	not yet started
Integrated Taxonomic Information System (ITIS)	Core Science Analytics & Synthesis	Core Science Systems	Guala	In Catalog	System-level Record	Yes	1
Ocean Biogeographic Information System (OBIS-USA)	Core Science Analytics & Synthesis	Core Science Systems	Mark Fornwall	In Catalog	System-level Record	Yes	1
	, , , , , , , , , , , , , , , , , , , ,		Gerald "Stinger"		,		
Biodiversity Information Serving Our Nation (BISON)	Core Science Analytics & Synthesis	Core Science Systems	Guala	In Catalog	System-level Record	Yes	1
Gap Analysis Program (GAP)	Core Science Analytics & Synthesis	Core Science Systems	Kevin Gergely	In Catalog	System-level Record	Yes	1
GAP Species Ranges	Core Science Analytics & Synthesis	Core Science Systems	Kevin Gergely	In Catalog	System-level Record	Yes	1
GAP Species Distribution Models GAP Land Cover Data	Core Science Analytics & Synthesis	Core Science Systems	Kevin Gergely	In Catalog	System-level Record	Yes	1
National Fish Habitat Partnership Data System (NFHP)	Core Science Analytics & Synthesis Core Science Analytics & Synthesis	Core Science Systems Core Science Systems	Kevin Gergely Andrea Ostroff	In Catalog In Catalog	System-level Record System-level Record	Yes Yes	1
Protected Areas Database of the United States (PADUS)	Core Science Analytics & Synthesis Core Science Analytics & Synthesis	Core Science Systems	Kevin Gergely	In Catalog	System-level Record	Yes	1
Multistate Aquatic Resources Information System (MARIS)	Core Science Analytics & Synthesis	Core Science Systems	Andrea Ostroff	In Catalog	System-level Record	Yes	1
USGS Photographic Library	USGS Libraries Program	Core Science Systems	Keith Van Cleave	In Catalog	System-level Record	Yes	1
USGS Publications Warehouse	USGS Libraries Program	Core Science Systems	Helen Tong	In Catalog	System-level Record	Yes	1
USA National Phenology Network (NPN)	Status & Trends Program	Ecosystems	Jake Weltzin	In Catalog	System-level Record	Yes	1
Geomagnetic Observatories	Geomagnetism Program	Natural Hazards	Duff Stewart	Not Yet Started	TBD	TBD	not yet started
Coastal & Marine Geoscience Data System (CMDGS)	Coastal & Marine Geology Program	Natural Hazards	Greg Miller	In Catalog	Individual Records	No	323
CoastalMap	Coastal & Marine Geology Program	Natural Hazards	John O'Malley	In Catalog	Individual Records	No	1167
InfoBank National Wetlands Research Center	Coastal & Marine Geology Program National Wetlands Research Center	Natural Hazards	Rex Saunders Cassie Thibodeaux	Not Yet Started In Catalog	Individual Records Individual Records	No No	not yet started 82
National Wetlands Research Center	National Wetlands Research Center	Ecosystems Energy & Minerals and Environmental	Cassie Inibodeaux	in Catalog	Individual Records	INO	02
Energy Data Finder	Central Energy Resources Team	Heath	Greg Gunther	In Catalog	Individual Records	No	505
Water NSDI Node	Water HQ	Water	Mike Ierardi	In Catalog	Individual Records	No	1101
South Florida Information Access (SOFIA)	Priority Ecosystem Science	Ecosystems	Heather Henkel	In Catalog	Individual Records	No	240
SageMap	Snake River Field Station, FRESC	Ecosystems	Tom Zarriello	Not Yet Started	Individual Records	No	not yet started
	Data Preservation, Informatics, and						
USGS Core Research Center Core Collection	Laboratories	Core Science Systems	Betty Adrian	In Catalog	System-level Record	No	1
LICCS Care Beautrels Contar Cuttings Collection	Data Preservation, Informatics, and Laboratories	Core Science Systems	Dotty Advisor	In Catalan	Custom lovel Decord	No	1
USGS Core Research Center Cuttings Collection	Data Preservation, Informatics, and	Core Science Systems	Betty Adrian	In Catalog	System-level Record	No	1
National Ice Core Laboratory	Laboratories	Core Science Systems	Betty Adrian	In Catalog	System-level Record	No	1
National Atlas	National Geospatial Program	Core Science Systems	Brian Fox ?	Not Yet Started	?	?	not yet started
							,

Contacts

Ben Wheeler: bwheeler@usgs.gov

Links:

SDC: data.usgs.gov

SDC dashboard: www.sciencebase.gov/catalog/dashboard

DOI CKAN: data.doi.gov

data.gov: data.gov

Geospatial Platform: geoplatform.gov

 USGS Science Data Catalog Team: Mike Frame, Viv Hutchison, Natalie Latysh, Lisa Zolly, Michelle Chang



